Hubble’s diagram and cosmic expansion

Robert P. Kirshner*

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

Contributed by Robert P. Kirshner, October 21, 2003

Edwin Hubble’s classic article on the expanding universe appeared in PNAS in 1929 [Hubble, E. P. (1929) Proc. Natl. Acad. Sci. USA 15, 168–173]. The chief result, that a galaxy’s distance is proportional to its redshift, is so well known and so deeply embedded into the language of astronomy through the Hubble diagram, the Hubble constant, Hubble’s Law, and the Hubble time, that the article itself is rarely referenced. Even though Hubble’s distances have a large systematic error, Hubble’s velocities come chiefly from Vesto Melvin Slipher, and the interpretation in terms of the de Sitter effect is out of the mainstream of modern cosmology, this article opened the way to investigation of the expanding, evolving, and accelerating universe that engages today’s burgeoning field of cosmology.

Today, ~70 years later, exquisite observations of the cosmic microwave background (2), measurement of light elements synthesized in the first few minutes of the universe (3), and modern versions of Hubble’s Law form a firm triangular foundation for modern cosmology. We now have confidence that a geometrically flat universe has been expanding for the past 14 billion yr, growing in contrast through the action of gravity from a hot and smooth Big Bang to the lumpy and varied universe of galaxies, stars, planets, and people we see around us. Observations have forced us to accept a dark and exotic universe that is ~30% dark matter with only 4% of the universe made of familiar protons and neutrons. Of that small fraction of familiar material, most is not visible. Like a dusting of snow on a mountain ridge, luminous matter reveals the presence of unseen objects.

Extensions of Hubble’s work with today’s technology have developed vast new arenas for exploration: extensive mapping using Hubble’s Law shows the arrangement of matter in the universe, and, by looking further back in time than Hubble could, we now see beyond the nearby linear expansion of Hubble’s Law to trace how cosmic expansion has changed over the vast span of time since the Big Bang. The big surprise is that recent observations show cosmic expansion has been speeding up over the last 5 billion yr. This acceleration suggests that the other 70% of the universe is composed of a “dark energy” whose properties we only dimly grasp but that must have a negative pressure to make cosmic expansion speed up over time (4–9). Future extension of the Hubble diagram to even larger distances and more precise distances where the effects of acceleration set in are the route to illuminating this mystery.

Hubble applied the fundamental discoveries of Henrietta Leavitt concerning bright Cepheid variable stars. Leavitt showed that Cepheids can be sorted in luminosity by observing their vibration periods: the slow ones are the intrinsically bright ones. By measuring the period of pulsation, an observer can determine the star’s intrinsic brightness. Then, measuring the apparent brightness supplies enough information to infer the distance.

This Perspective is published as part of a series highlighting landmark papers published in PNAS. Read more about this classic PNAS article online at www.pnas.org/misc/classics.shtml.

*E-mail: kirshner@cfa.harvard.edu.

1There are just 73 citations of Hubble’s original paper in NASA’s Astrophysics Data System. There are 1,001 citations of ref. 7.

© 2003 by The National Academy of Sciences of the USA
Hubble used the 100-inch Hooker Telescope at Mount Wilson to search for these “standard candles” and found Cepheids in the fuzzy Andromeda Nebula, M31. From the faint appearance of those Cepheids, Hubble deduced that M31 and the other “extra-galactic nebulae” are not part of our own Milky Way galaxy, but “island universes” equivalent to the Milky Way: vast systems of billions of stars separated from one another by millions of light years. This finding was made in 1924, and if he had done nothing more than to show that the Milky Way is not the universe, Hubble would have been an important figure in the history of astronomy. But 5 yr later in his PNAS article, Hubble was able to show something even more astonishing by plotting the velocities of galaxies against their distances.

Reading Hubble’s article is a healthy reminder of how much clearer things become with 70 yr of hindsight. For example, although the Cepheids lie at the foundation of Hubble’s distance scale, the distances to most of the objects in his 1929 article were not determined by Cepheids themselves, but by the brightest stars in galaxies or by the luminosity of the galaxies themselves. In recent years, by using the superb resolution of the Hubble Space Telescope, named in Edwin Hubble’s honor, it is finally possible to measure individual Cepheids in galaxies in the Virgo cluster that are the most distant entries in Hubble’s original table of galaxy redshifts and distances (10, 11). The quantitative agreement of modern measurements with Hubble’s original distance scale is not good! Modern distances to the same galaxies, reckoned to be accurate to 10%, are seven times larger than the distances Hubble plots horizontally in Fig. 1. Hubble’s essential contribution was a consistent set of distances to galaxies that allowed him to glimpse the underlying relation between distance and velocity. Although his distances had serious errors due to confusing two types of Cepheids, and blurring bright gas clouds with bright stars, in 1929, Hubble was able to sort nearby galaxies from distant ones well enough not to miss the connection between distance and velocity.

The other axis of the Hubble diagram (subtly mislabeled in the original) shows not only that we live in a spacious universe populated by billions of galaxies like the Milky Way, but also that the galaxies are embedded in an expanding fabric of space and time. The Hubble diagram plots velocity against distance. Astronomers measure the velocity of a galaxy from its spectrum by taking the light from a galaxy’s image at the focus of a telescope and passing it through a slit and a prism to create a dispersed rainbow, subtly marked by dark lines. These absorption lines are produced by atoms in the atmospheres of stars. Atoms absorb light at specific wavelengths, matching the energy jumps for electron orbits dictated by quantum mechanics. Radial velocities show up as shifts in the wavelengths of the lines from the galaxy compared with the spectra of the same atoms at rest in the observatory: blue-shifts for objects approaching us and redshifts for objects receding. The fractional shift of the wavelength, \( \Delta \lambda / \lambda = 1 + z \), where \( z \) is the redshift. This result can be expressed as a velocity, \( c \), where \( c \) is the speed of light, 300,000 km/s.

The program of measuring galaxy spectra had been initiated a decade earlier by Vesto Melvin Slipher at the Lowell Observatory in Arizona. By 1923, after heroic efforts with small telescopes and slow spectrographs, Slipher had compiled a list of velocities for 41 galaxies, 36 of which were receding from us, and the largest of which was moving away at 1,800 km/s. This intriguing list was published in Arthur Stanley Eddington’s textbook on general relativity, The Mathematical Theory of Relativity. Hubble cites no source for the radial velocities in table 1 of ref. 1, except for the four new ones from his Mount Wilson colleague, Milton Humason, but every one of the galaxies is one of Slipher’s, and the list of velocities is almost identical to that in Eddington’s book. Hubble’s original contribution in 1929 was to grasp the connection of distance with velocity, and his subsequent effort was to pursue the consequences of this amazing fact. Hubble and Humason poured prodigious effort into measuring redshifts at the 100-inch, rapidly expanding the reach of the Hubble diagram beyond the 1,000 km/s velocities shown in Fig. 1. Although Slipher had begun the field of galaxy spectra a decade earlier and measured the velocities that Hubble used in his 1929 article, Hubble soon became the towering figure in exploring the realm of the nebulae.

The connection between general relativity and cosmic velocities was lurking in the background of Hubble’s work. In 1917, Einstein had shown how to construct a universe that was static by introducing a “cosmological constant” into his equations. This matched well with the idea, current before Hubble’s 1924 measurement of the distances to the nebulae, of a small and static “universe” that was confined to the stars of the Milky Way galaxy. In Leiden, the eminent de Sitter had shown that there was another, formally static, solution to Einstein’s equations in which particles would scatter with an acceleration increasing with distance and signals sent from one observer to another would show a redshift. More physical solutions to Einstein’s equations, constructed for an expanding universe, were worked out by Friedmann in 1922. But those were not the models Hubble was thinking of when he plotted his data. Hubble was looking for the de Sitter effect.

Putting distances and velocities together on the graph shown as Fig. 1 in Hubble’s classic article, anybody can see that the velocity is more or less proportional to the distance. What transforms this bland diagram into a profound discovery is an understanding that the pattern Hubble found is exactly what you would expect for any observer in a universe expanding in all directions. Hubble’s diagram does not imply that we are at the center of the universe, but it does show that the universe is dynamic, definitely not static, as Einstein had assumed in 1917.

In the text of his article, Hubble says, “the outstanding feature is the possibility that the velocity-distance relation may represent the de Sitter effect.” It is probably not a coincidence that Hubble was looking for this effect: he was in Leiden in 1928 for a conference on galaxies, and he had the opportunity to talk with de Sitter. Hubble notes that one aspect of the de Sitter world model is an apparent “acceleration,” and he makes the plausible assumption that the “linear relation found in the present discussion is a first approximation representing a restricted range in distance.” This particular aspect of Hubble’s article has seemed quaint and puzzling.

It has seemed quaint because, from 1931 to about 1995, almost nobody was talking about cosmic acceleration. As a result of Hubble’s own work, even Einstein and de Sitter stopped talking about their old models with a cosmological constant, and the observational focus shifted to finding the numerical value of the Hubble constant and to measuring the gravitational effect of cosmic deceleration in an expanding Friedmann model. And it seemed puzzling because, in modern parlance, an Einstein–de Sitter model is an expanding Friedmann model with a flat geometry, and the original de Sitter effect is a historical curiosity. But today, in light of recent work that suggests that we do live in an accelerating universe of the Einstein–de Sitter type, with Euclidean space, Hubble’s illusion to acceleration seems oddly, perhaps falsely, prescient. The slope of the line in the Hubble diagram is called the Hubble constant \((H_0)\), directly related to the age of the
universe: Hubble’s Law says that velocity = \( H_0 \times \) distance and, because time = distance/velocity, there is a natural Hubble time, \( t_0 \), associated with the Hubble expansion, \( t_0 = \frac{\text{distance}}{H_0 \times \text{distance}} = \frac{1}{H_0} \).

Nearby objects recede slowly, and more distant ones recede rapidly, but both would take the same time to get where they are in a universe that expands at a constant rate, and that time is given by \( 1/H_0 \). So the Hubble constant sets the time scale from the Big Bang to today.4

Although Hubble’s 1929 distances were too small by a factor of 7, his conclusion about the nature of cosmic expansion was still valid because all his distances were too small by about the same factor. The form of the relation, velocity proportional to distance, is not changed by this scale error, although the numerical values for the distances, and for the Hubble constant (which Hubble modestly called K) is far from the modern value. In this classic article, Hubble quotes values of K of 530 and 500 km/s/megaparsec. Staring at his original Hubble diagram, you can see that there is a handful of nearby galaxies with blueshifts, and a large scatter of velocities at any given distance. Hubble shrewdly used plausible methods to average the data for galaxies that are at the same distance to make his result stand out more clearly from the noise. He was fortunate to have data that behaved so well.

Over time, improved understanding of the stars being used, the role of absorption by dust, and the local calibration of the distance scale led to large revisions in the cosmic distance scale, the Hubble constant, and in the inferred Hubble time. In Hubble’s time, \( t_0 \) was \( \sim 2 \) billion yr, which was already in conflict with the larger age of the Earth inferred from radioactive decay. The Earth should not be older than the universe in which it formed. This conflict with the age of the Earth and a similar problem with the ages of the stars was a chronic embarrassment during the decades when the Hubble constant was poorly known. The disagreement made it difficult to accept the reality of cosmic expansion acting over cosmic time, and Hubble was always quite circumspect on the interpretation of his discovery. But, as shown in Fig. 2 of this Perspective, my colleague John Huchra’s compilation of the numerical value of the Hubble constant shows how the prevailing value has been dropping over the decades. The quoted error bars are chronically much smaller than the drift in the mean value over time. The systematic errors are always underestimated. This plot lends weight to the aphorism that astrophysicists are always wrong, but never in doubt.

Modern work, closely tied to the Cepheids in Virgo cluster galaxies observed with the Hubble Space Telescope gives \( H_0 = 72 \pm 2 \pm 7 \) km/s/megaparsec (9). The errors quoted are one sigma, with the first being the statistical error, and the second, larger error being the systematic uncertainty due to factors like the chemical composition of the Cepheids in different galaxies, the distance to the Large Magellanic Cloud to which the distance comparison is made, and the calibration of the camera on the Hubble Space Telescope. As in the past, we believe these error bars are correct (although for a contrasting view, see ref. 10). But now, the convergence from completely independent methods such as time delays in gravitational lenses, scattering of microwave background photons by hot gas in galaxy clusters, and the physics of supernova atmospheres is beginning to be significant (12–16). With independent paths, systematic errors can be exposed. We are, at least, coming to the end of the search for the Hubble constant.

The remarkable result of this long path of revision is that the Hubble time is now taken seriously. The age of the universe implied by the modern Hubble constant with constant expansion is \( \sim 14 \) billion yr. This result is in good accord with the theoretical ages of stars. The oldest stars in our galaxy have ages, based on computations of stellar evolution through nuclear burning, of \( \sim 12.5 \pm 1.5 \) billion yr, just enough younger than the Hubble time to fit comfortably into a scheme where galaxies form promptly after the Big Bang (17). Even with the added wrinkle of cosmic deceleration and cosmic acceleration, the best value from the Hubble diagram for the elapsed time since the Big Bang is \( \sim 13.6 \pm 1.5 \) billion yr (18). The expansion is no illusion; it is cosmic history.

As in Hubble’s original article, where he used the very brightest stars and the light from entire galaxies, the modern path to deeper distance measurements is through a brighter standard candle than the Cepheids. Before Hubble, astronomers had, from time to time, noted new stars that flared up in extragalactic nebulae like M31 and its cousins. In our own galaxy, these new stars are called “novae.” Once Hubble had established that the distances to these nebulae were millions of light years, the true nature of these novae became clear. Because they were at distances a thousand times larger than novae in the Milky Way, they must be a million times more energetic. Exploding stars in other galaxies were dubbed “supernovae” by Fritz Zwicky, Hubble’s contemporary down Lake Avenue in Pasadena at the California Institute of Technology. The light output of one particular type of supernova is \( \sim 4 \) billion times that of the sun. These “type Ia” supernovae can be seen half way across the visible universe, and, even better, they have a fairly narrow distribution in intrinsic brightness. As a

---

4The conventional units of the Hubble constant are a bit obscure: 1 megaparsec (Mpc) = \( 10^6 \) parsec = \( 3.26 \times 10^9 \) m. A Hubble constant of 70 km/s/Mpc corresponds to \( 2.27 \times 10^{-18} \) s⁻¹. Then, the Hubble time is \( 1.227 \times 10^{18} \) s or \( 13.9 \times 10^9 \) yr.
result, they make good distance indicators. Refined methods for analyzing the observations of type Ia supernovae give the distance to a single event to better than 10% (19, 20). The best modern Hubble diagram, based on well observed type Ia supernovae out to a modest distance of \( \sim 2 \) billion light years, is shown in Fig. 3, where the axes are chosen to match those of Hubble’s original linear diagram (to mask our uncertainties, astronomers generally use a log-log form of this plot as in Fig. 4). Far beyond Hubble’s original sample, Hubble’s Law holds true.

In table 2 of his original article (1) (reproduced as Table 1, which is published as supporting information on the PNAS web site), Hubble inverted the velocity–distance relation to estimate the distances to galaxies of known redshift. For galaxies like NGC 7619 for which he had only Humason’s shift. For galaxies like NGC 7619 and the distances to galaxies of known redshift (reproduced as Table 1, which is published as supporting information on the PNAS web site), Hubble inverted the velocity–distance relation to estimate the distances to galaxies of known redshift. For galaxies like NGC 7619 for which he had only Humason’s shift.

Part of the difficulty with the interpretation came from alternative views, notably by the local iconoclast, Fritz Zwicky, who promptly sent a note to PNAS in August 1929 that advocated thinking of the redshift as the result of an interaction between photons and intervening matter rather than cosmic expansion (26). The reality of cosmic expansion and the end of “tired light” has only recently been verified in a convincing way.

While the nature of the redshift was a bubbling discussion in Pasadena, Olin Wilson of the Mount Wilson Observatory staff suggested that measuring the time it took a supernova to rise and fall in brightness would show whether the expansion was real. Real expansion would stretch the characteristic time, about a month, by an amount determined by the redshift (27).

This time dilation was sought in 1974, but the sample was too small, too nearby, and too inhomogeneous to see anything real (28). It was only with large carefully measured and distant samples of SN Ia (29, 30) and more thorough characterization of the way supernova light curves and supernova luminosities are intertwined (31, 32) that this topic
Supernova explosions behave better. As discrete physical events with a well-defined energy, supernovae of type Ia work well as standard candles over a very large range in redshift. These explosions allow us to look back to the time when the universe was young and to see the effects of changes in the rate of expansion reflected in the Hubble diagram. During the first stages of this work, the observers expected to see the deceleration that mass would cause (39). The first reports from the Supernova Cosmology Project using supernovae confirmed that view (40). However, better data sets for the Hubble diagram of distant supernovae from the High-Z Supernova Team and from the Supernova Cosmology Project (7, 8) showed the surprising result that the expansion of the universe has been speeding up during the 5-billion-yr interval while the light from a distant supernova has been in flight to our telescopes.

The most recent summary of the supernova data by Tonry et al. (18) shows that this result is robust and fits well with recent results on the microwave background and large-scale galaxy distributions. The acceleration is attributed to the negative pressure of a smoothly pervasive component of the universe: the dark energy (4). One possibility is that the dark energy is the cosmological constant looked at another way: as a vacuum energy, rather than as a curvature term in Einstein’s equations. If this picture is really right,

could be explored with confidence. Best of all was to have supernovae at high redshift where the effect would not be subtle. Bruno Leibundgut et al. (33) showed that the light curve for one object, SN 1995K, at a redshift of $z = 0.479$ matched the light curve of a nearby SN Ia, but only when stretched by time dilation by a factor of $1 + z$. Similarly, the time evolution of the spectrum for the type Ia supernova SN 1996J by $z = 0.574$ was also stretched out by the redshift (34). Goldhaber et al. (32) examined the effect of time dilation by using a large set of high-$z$ supernovae and found results in complete accord with the expectations of real cosmic expansion, not photon fatigue. A second prediction of the expansion idea is that the surface brightness of a galaxy should decrease as $(1 + z)^4$. This “Tolman dimming” has finally been observed by Rubin and Sandage (35). The idea of tired light has now been put to rest.

Einstein’s idea of a static universe, suspended between gravity pulling inward and the cosmological constant making the universe expand, was ruled out by Hubble’s data. Legend has it that Einstein, much later, referred to the cosmological constant as his “greatest blunder” (36). In 1947, Einstein wrote, “Since I introduced this term, I had always a bad conscience. . . . I am unable to believe that such an ugly thing is actually realized in nature” (37). In their 1932 farewell to the cosmological constant (also published in PNAS), Einstein and de Sitter were more measured: “an increase in the precision of the data derived from observations will enable us in the future to fix its sign and to determine its value” (38).

The cosmological constant was banned by Einstein’s curse from serious cosmological discussion from 1932 to about 1995. The observational program of practical cosmology shifted to measuring two parameters: the Hubble constant and the deceleration that gravity produces over time. The goal was to construct a Hubble diagram in which the most distant objects were sufficiently far to show a clear deviation from the linear law seen by Hubble in 1929. In 1989, as in 1929, the problem was not with the redshifts, but with the distances. Using galaxies to measure distances proved frustrating: the stars that make up galaxies fade as galaxies age, but galaxies accrete more stars, and it was too hard to tell whether distant, young galaxies were intrinsically brighter or dimmer than their counterparts nearby.

![Fig. 5. Large scale structure inferred from galaxy redshifts. Each dot in this plot marks a galaxy whose distance is estimated from its redshift by using Hubble’s Law. From the 2DF Galaxy Redshift Survey (24).](image)

![Fig. 6. Deviations in the Hubble diagram. Each point in this plot shows the difference at each redshift between the measured apparent brightness and the expected location in the Hubble diagram in a universe that is expanding without any acceleration or deceleration. The blue points correspond to median values in eight redshift bins. The upward bulge at $z = 0.5$ is the signature of cosmic acceleration. The hint of a turnover in the data at the highest redshifts, near $z = 1$, suggests that we may be seeing past the era of acceleration driven by dark energy back to the era of deceleration dominated by dark matter. From top to bottom, the plotted lines correspond to the favored solution, with 30% dark matter and 70% dark energy, the observed amount of dark matter (30%) but no dark energy, and a universe with 100% dark matter (from ref. 18).](image)
then constructing a precise Hubble diagram in the era where the acceleration has its onset, and pushing the Hubble diagram for type Ia supernovae to redshifts beyond 1 will help to pin down the nature of the dark energy.

This work is already underway. Fittingly, this extension of Hubble’s work is being carried out using the Hubble Space Telescope as well as ground-based observatories. One very high redshift supernova was discovered in 1996 (41) at redshift 1.7, well over halfway back to the Big Bang, and many more will be found with the new Advanced Camera for Surveys that was installed on the Hubble Space Telescope by shuttle astronauts a year ago (42, 43). If the dark energy picture is right, we should expect to see back past the era of acceleration, to an earlier era of deceleration when dark matter ruled the dynamics of the universe as hinted at by Fig. 6. Hubble hoped to understand cosmic expansion by seeing the higher-order terms that lay beyond the linear expansion of the nearby sample; we are now looking deep into the past at the limits of today’s technology to observe directly these changes in cosmic expansion. As Hubble said in The Realm of the Nebulae, “We measure shadows, and we search among ghostly errors of measurement for landmarks that are scarcely more substantial. The search will continue” (44). Hubble’s article had velocities from Trumpler without citation, distances wrong by a factor of seven, reference to de Sitter’s strange kinematic model, and was not enough to convince Hubble himself of the reality of cosmic expansion, but that article in PNAS pointed the way to understanding the history of the universe, and the continuing search among the “ghostly errors of measurement” has led to a deeply surprising synthesis of dark matter and dark energy.

I am very grateful to John Huchra, Saurabh Jha, Brian Schmidt, Tom Matheson, and John Tonry for help with the text and illustrations and to the entire High-Z Supernova Search Team for their hard work in building on Hubble’s foundation. Work on supernovae and the Hubble constant at Harvard University is supported by National Science Foundation Grants AST 02-05808 and AST 02-06329 and by the National Aeronautics and Space Administration through Space Telescope Science Institute Grants GO-08641 and GO-09118.